



Aligning science assessment standards: Louisiana and the 2009 National Assessment of Educational Progress (NAEP)



Institute of Education Sciences
U.S. Department of Education



Aligning science assessment standards: Louisiana and the 2009 National Assessment of Educational Progress (NAEP)

July 2007

Prepared by

**Michael Timms
WestEd**

**Steven Schneider
WestEd**

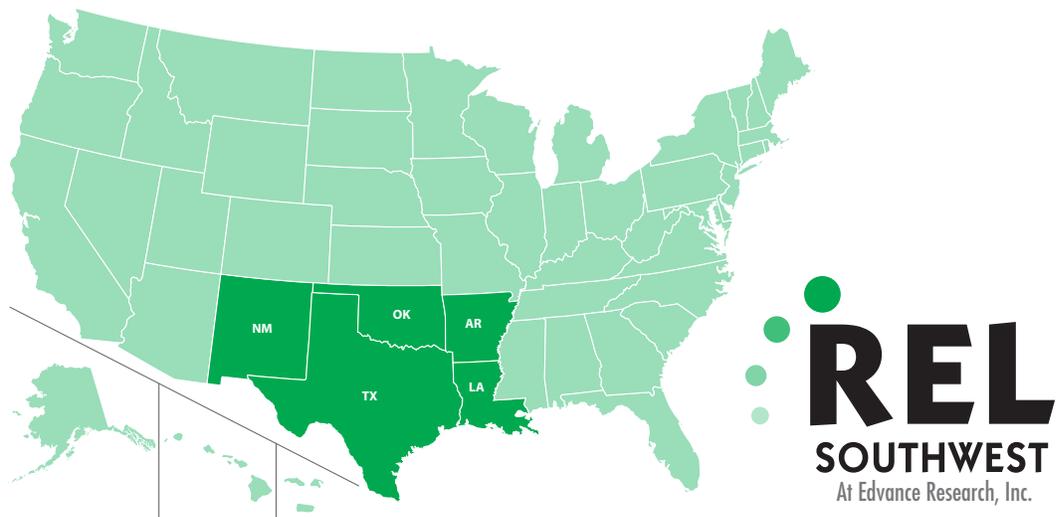
**Cindy Lee
WestEd**

**Eric Rolfhus
REL Southwest**



Institute of Education Sciences

U.S. Department of Education



Issues & Answers is an ongoing series of reports from short-term Fast Response Projects conducted by the regional educational laboratories on current education issues of importance at local, state, and regional levels. Fast Response Project topics change to reflect new issues, as identified through lab outreach and requests for assistance from policymakers and educators at state and local levels and from communities, businesses, parents, families, and youth. All Issues & Answers reports meet Institute of Education Sciences standards for scientifically valid research.

July 2007

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-06-CO-0017 by Regional Educational Laboratory Southwest administered by Edvance Research. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This report is in the public domain. While permission to reprint this publication is not necessary, it should be cited as:

Timms, M., Schneider, S., Lee, C., & Rolffhus, E. (2007). *Aligning science assessment standards: Louisiana and the 2009 National Assessment of Educational Progress (NAEP)* (Issues & Answers Report, REL 2007–No. 020). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>.

This report is available on the regional educational laboratory web site at <http://ies.ed.gov/ncee/edlabs>.

Summary

Aligning science assessment standards: Louisiana and the 2009 National Assessment of Educational Progress (NAEP)

This policy research document is intended for Louisiana policymakers to use when examining possible changes to the state assessment’s alignment with the National Assessment of Educational Progress (NAEP). The 2009 NAEP test is not yet in existence, so the purpose of this report is to give policymakers a headstart in determining where they might, if they so decide, begin to make changes in their assessment standards and specifications to develop an assessment system more closely aligned to that used for the NAEP.

Reviewers observed that NAEP standards are based on descriptions of phenomena, whereas Louisiana’s standards are based on explanations of what a student does to explore the phenomena. They also found that NAEP standards tend to be widely inclusive (including, for example, heat and electrical conductivity in the same standard) whereas most states tend to see such topics as part of two separate standards or benchmarks. Reviewers found Louisiana’s science standards to be rigorous and found consistent spiraling through the grade levels. However, they also noted that teachers might more easily discern nuances in the standards if the grade level expectations, as well as the “key concepts” within the assessment guides, were integrated with the benchmarks.

In comparing Louisiana benchmarks and grade level expectations with the NAEP, the overall alignment ratings for elementary, middle, and high school are generally very high. The overall alignment rating for Louisiana science content and NAEP grade 4 is 2.6. (A rating of 1 indicates no alignment and a rating of 3, full alignment.) For grade 8, the alignment rating is 2.1. For grade 12, the rating is 2.5. Louisiana middle school content is partially aligned with the NAEP grade 8, while grade 4 and high school are most often fully aligned.

The instances of partial alignment at all grade levels are due primarily to the fact that Louisiana’s standards often imply NAEP content and the NAEP is often more detailed in its presentation of content. In addition, Louisiana contains many benchmarks that are unaddressed by the NAEP content statements. Generally, the combination of Louisiana’s benchmarks and grade level expectations at all grade levels aligns very well with the NAEP content statements, because the grade level expectations often parallel NAEP statements in their level of detail.

This report reveals current alignment issues between the state’s tests and the future NAEP tests and may be especially important to policymakers considering revising

science standards and assessments in line with No Child Left Behind requirements for state science tests in elementary, middle, and high schools. If state policymakers wish to increase the alignment between the state assessments and the NAEP, an area to consider is developing the comprehensive science task into a hands-on performance task. Revising assessments requires considerable time and resources, so policymakers must carefully consider their capacity to make changes and the degree to which such changes will benefit students.

Grade 4 alignment

All grade 4 NAEP content statements are at least partially addressed by Louisiana benchmarks and grade level expectations. The Louisiana benchmarks also contain many items that are not listed in the NAEP's content statements. The combination of Louisiana's benchmarks and grade level expectations aligned very well with the NAEP content statements. Many partially aligned content items were found to be so because they imply content explicitly stated by the NAEP.

Reviewers observed that NAEP standards are based on descriptions of phenomena, whereas Louisiana's standards are based on explanations of what a student does to explore the phenomena. They also found that NAEP standards tend to be widely inclusive (including, for example, heat and electrical conductivity in the same standard) whereas most states tend to see such topics as part of two separate standards or benchmarks. Reviewers found Louisiana's science standards to be rigorous and found consistent spiraling through the grade levels. However, they also noted that

teachers might more easily discern nuances in the standards if the grade level expectations, as well as the "key concepts" within the assessment guides, were integrated with the benchmarks.

The majority of NAEP content statements are fully addressed by various combinations of Louisiana benchmarks and grade level expectations, and the overall alignment rating for Louisiana and NAEP science content at grade 4 is 2.6.

Grade 8 alignment

The Louisiana benchmarks cover many content topics not assessed by the NAEP, while most of the Louisiana content in its benchmarks and grade level expectations only partially addresses the NAEP content statements. This is primarily because Louisiana's standards often imply NAEP content and because the NAEP is often more detailed in its presentation of content. The overall alignment rating for Louisiana and the NAEP grade 8 is 2.1, indicating partial alignment.

Grade 12 alignment

The Louisiana benchmarks contain many items not assessed by the NAEP. Several Louisiana benchmarks were noted for alignment to distinct parts of a more detailed NAEP content statement. Overall, Louisiana is well aligned to the NAEP. Many of the Louisiana high school standards only partially address the NAEP content statements at grade 12, because Louisiana's benchmarks often imply NAEP content and because the NAEP is often more detailed in its presentation of content. The overall alignment rating is 2.5.

Test specifications

The Louisiana assessment guides, which define the specifications for the state tests, ensure that testing of student knowledge and skills does not rely solely on multiple-choice items by including short constructed-response items and a comprehensive science task at each grade level. That enables a wider range of knowledge types to be tested than can be tested with multiple-choice alone. Louisiana records its proportions differently than the NAEP, so it is difficult to directly compare the relative amounts of testing time devoted to different topics. However, when focusing just on the three topic areas tested by the NAEP, the proportions of NAEP testing times are the

same as the proportions of points in the Louisiana test at grade 4 and in high school, and they are similar at grade 8, where Louisiana's points are allocated equally across subjects while NAEP students are tested slightly more on Earth and space science. Overall, there is a match between the test specifications in Louisiana's assessment guides and the NAEP science assessment and item specifications.

Standards and test specifications represent the starting point for the development of tests and test items. In the ideal alignment study state science assessments would be compared with NAEP assessments directly at the item level. At some future date the NAEP 2009 assessment items may be available for such a study.

July 2007

TABLE OF CONTENTS

Summary	iii
Background to the study	1
Content alignment at grade 4	3
Areas of full alignment	4
Areas of partial alignment	4
Areas of nonalignment	5
Areas where Louisiana benchmarks go beyond the NAEP content statements	5
Summary of grade 4 alignment	5
Content alignment at grade 8	6
Areas of full alignment	7
Areas of partial alignment	7
Areas of nonalignment	7
Areas where Louisiana benchmarks go beyond the NAEP content statements	7
Summary of grade 8 alignment	8
Content alignment at grade 12	8
Areas of full alignment	8
Areas of partial alignment	8
Areas of nonalignment	9
Areas where Louisiana benchmarks go beyond the NAEP content statements	10
Summary of NAEP grade 12 alignment	10
Test specifications alignment	10
Appendix A The documents compared	15
Appendix B How the study was conducted	19
Appendix C Content alignment for grade 4	25
Appendix D Content alignment for grade 8	47
Appendix E Content alignment for grade 12	60
References	83
Box 1 Methodology	3
Figures	
1	The majority of Louisiana elementary school standards fully address National Assessment of Educational Progress content statements 4
2	The majority of Louisiana middle school standards partially address National Assessment of Educational Progress content statements 6
3	The majority of Louisiana high school standards fully address National Assessment of Educational Progress content statements 9
B1	Crosswalk instrument 22

Tables

1	Average ratings of alignment of Louisiana grade 5 benchmarks and grade level expectations and National Assessment of Educational Progress grade 4 science content statements	4
2	Average ratings of alignment of Louisiana middle school science benchmarks and grade level expectations and National Assessment of Educational Progress grade 8 science content statements	6
3	Average ratings of alignment of Louisiana high school benchmarks and grade level expectations and National Assessment of Educational Progress grade 12 science content statements	9
4	Proportions of different item types on the Louisiana science assessment	12
5	Approximate testing time allocated to different science topics on the Louisiana science assessment	13
6	Comparison of the proportions of testing time allocated to the National Assessment of Educational Progress science topics	13
A1	National Assessment of Educational Progress distribution of items and standards by content area and grade	15
A2	Number of multiple-choice, short-answer, and extended constructed-response items by strand—Louisiana grades 4 and 8	16
A3	Number of multiple-choice, short-answer, and extended constructed-response items by strand—Louisiana Graduation Exit Examination grade 11	17
A4	Distribution of benchmarks and grade level expectations by strand—grades 4 and 8	18
A5	Distribution of benchmarks and grade level expectations by strand—Graduation Exit Examination and high school grade level expectations	18
C1	Alignment of National Assessment of Educational Progress grade 4 science and Louisiana grade 4 benchmarks and grade level expectations	25
C2	Louisiana grade 4 standards not covered by National Assessment of Educational Progress grade 4 content	46
D1	Alignment of National Assessment of Educational Progress grade 8 science and Louisiana grade 8 benchmarks and grade level expectations	47
D2	Louisiana grade 8 standards not covered by National Assessment of Educational Progress grade 8 content	59
E1	Alignment of National Assessment of Educational Progress grade 12 science and Louisiana grade 11 benchmarks and grade level expectations	60
E2	Louisiana grade 12 high school benchmarks not covered by National Assessment of Educational Progress grade 12 content	82

This policy research document is intended for policymakers to use when examining possible changes to the state assessment's alignment with the National Assessment of Educational Progress (NAEP).

BACKGROUND TO THE STUDY

This report presents the findings of an alignment study comparing the new science framework for the 2009 NAEP and the accompanying science assessment and item specifications with the Louisiana state science assessment. More details about the documents compared are in appendix A. The study was conducted for the Regional Education Laboratory Southwest, funded by the Institute of Education Sciences, to provide research and support to Arkansas, Louisiana, New Mexico, Oklahoma, and Texas. The study was undertaken in anticipation of a growing need in the region to be better informed about how state assessment

standards in science compare with those tested in the NAEP.

The 2009 NAEP test is not yet in existence, so the purpose of this report is to give policymakers a headstart in determining where they might, if they so decide, begin to make changes in their assessment standards and specifications to develop an assessment system more closely aligned with that used for the NAEP.

Five factors make this study timely. First, the importance of state science assessments has been increased by the No Child Left Behind Act of 2001. Beginning in the 2007/08 school year, states are required to administer science assessments to all students in each of the elementary, middle, and high school levels, holding states and local school districts accountable for student academic achievement in science (NCLB, 2001).

Second, the NAEP is increasingly being used as a benchmark against which student achievement across the nation can be compared (Linn, 2005; Linn, Baker, & Herman, 2005). The NAEP has been dubbed the “nation’s report card,” and when fresh NAEP results are released—as they were for science in 2006, following an administration of the test in 2005—the media report the results (Cavanagh, 2006a, 2006b). Although states are not sanctioned for failing to demonstrate NAEP student performance improvement, NAEP data do provide an external accountability benchmark and serve to verify student achievement on state assessments. In fact, the National Center for Education Statistics has a website (<http://nces.ed.gov/nationsreportcard/nde/statecomp/>) that allows anyone to create customized comparative reports based on the latest NAEP data. So anyone can create tables that compare states and jurisdictions based on the average scale scores for selected groups of public school students within a single assessment year, or compare the change in performance between two assessment years.

Third, NAEP data are being used more in education research to investigate how the No Child Left

Behind Act provisions have played out in different states. For example, Olson (2005) compared the percentages of students at or above the proficient level on the 2005 state grade 8 mathematics assessments in 33 states. The study showed that, on average, 33 percent more students scored at or above the proficient level according to the state assessments than did so according to the NAEP. As yet, no similar study has been done of science, but with the release of the 2005 NAEP results it is now possible to do so.

Fourth, political attention is beginning to focus on using the NAEP as a yardstick for measuring state standards (Olson, 2007). In January 2007 two bills were introduced in Congress, one seeking to encourage states to benchmark their own standards and tests to the NAEP and the other calling for states to adopt voluntary “American education content standards” in mathematics and science that would be developed by the National Assessment Governing Board, the body responsible for the NAEP. These issues will doubtless be topics of debate in the upcoming reauthorization of the No Child Left Behind Act.

Fifth, the standards and test specifications that form the blueprint for the content the NAEP science assessment covers and the types of items it uses were revised in 2006. The 2009 NAEP framework takes account of the latest knowledge on science learning and assessment, which suggests that measuring student understanding involves much more than assessing factual knowledge. It defines the science knowledge and skills that science-literate students should possess at grades 4, 8, and 12. The assessment itself, while retaining some familiar paper-and-pencil assessment

formats, will also include student performance assessments in both classroom settings and computer simulations. The 2009 NAEP framework will determine the shape of NAEP science assessments through 2017, setting the direction of science assessment across the nation.

These factors are working together to gradually raise the NAEP to a de facto national benchmark, and states naturally want to know how well their state standards align with the NAEP so they can make informed decisions about possible changes to their own standards and assessment systems. This report describes the results of a systematic alignment study of science assessment standards conducted for that purpose. Details of the study are in appendix B.

The intent of this report is to inform those in the Louisiana Department of Education responsible for shaping the state assessment in science how the current assessment standards and test specifications compare with those of the NAEP 2009 assessment.

Similar reports have been completed for Arkansas, New Mexico, Oklahoma, and Texas, but there is no intent to compare Louisiana with these states. This report shows where there is good content alignment with NAEP standards, identifies where there is partial alignment, pinpoints NAEP standards where there are no corresponding state standards, and highlights where the Louisiana standards go beyond the NAEP. It also deals with the assessment specifications, showing what percentages of the NAEP assessment at each grade level are devoted to different science topics and comparing that to the coverage of the topics in the Louisiana assessment. And it compares the proportions of types of items used to test students’ science knowledge and skills. Through comprehensive comparative analysis, the report provides a way for the Louisiana Department of Education to gauge how well its tests are doing in covering the depth of science understanding expected on the NAEP.

The NAEP science standards were compared with the Louisiana elementary level benchmarks and the Louisiana grade level expectations. Louisiana recommended the use of its grade level expectations for this study, as the state’s assessment specialists indicated that assessment content was drawn not only from the benchmarks, but also from the grade level expectations. Thus, while the

Several factors are working to raise the NAEP to a de facto national benchmark, and states want to know how well their state standards align with it

alignment studies performed for other states in the Southwest Region were completed using only the curriculum content or grade level expectations in the assessment frameworks, Louisiana's alignment was performed using the complete set of its benchmarks and grade level expectations.

The results are presented in the summary tables and narratives in the sections that follow. Those sections provide an analysis that highlights the differences found between NAEP's content and Louisiana's content as presented by the Louisiana Educational Assessment Program assessment guides (grades 4, 8, and 11) and the Louisiana grade level expectations. For more detail about the alignment of the state content to the individual content statements of the NAEP, turn to the tables in appendixes C–E. They show exactly which Louisiana

standards align with a particular NAEP statement and, in cases of partial alignment, explain why the alignment is incomplete. For a discussion of methodology, see box 1 and appendix B.

CONTENT ALIGNMENT AT GRADE 4

For grade 4, the NAEP provides 33 distinct content statements (displayed in parentheses in table 1). Twenty-one of these content statements (64 percent) are fully addressed by Louisiana content in the benchmarks and grade level expectations, and 12 (36 percent) are partially addressed. No NAEP content statements are unaddressed.

The average alignment rating for grade 4 is 2.6 (table 1). The majority of content statements were

BOX 1

Methodology

The chief research questions driving this study were: “To what extent do current state assessment standards cover the content on which NAEP 2009 assessments will be based?” and “To what extent do current state assessment specifications align with the NAEP 2009 assessment specifications?”

The methodology used to answer these questions followed the successful pattern of a similar study conducted by WestEd in New England, which examined the alignment of math and reading standards with the NAEP. The methodology developed by WestEd for the New England study was designed to include all the most prominent alignment methodologies, which are discussed in appendix B. Thus far, alignment studies and methods have focused on aligning standards with tests, whereas the objective

of this study was to compare one set of assessment standards and specifications with another. The methodology in this study, however, is based on methods for aligning standards with tests, because similar principles are used in both types of alignments.

In this study reviewers followed the methodology of the portion of the previous study examining alignment between two sets of standards. Test blueprints were examined to find correspondence between the two documents. Reviewers performed gap analyses to identify content included in one set of standards but not the other, identified issues of order so they could reveal differences in the grade levels at which standards appear, and examined the degree to which the standards and assessments cover content to the same depth and have similar cognitive demands (depth-of-knowledge consistency) and the degree to which assessments cover the same range of content as

the corresponding standards (range-of-knowledge correspondence) to determine whether there is a match between the state and the NAEP in the level of detail, cognitive demands, and range of content covered. A coding scheme was used to indicate alignment issues and reviewer ratings, and a matrix-like format was created to facilitate alignment.

Reviewers attended several training sessions, conducted individual reviews, and then met in teams of two to reach consensus on ratings. This consensus method was designed to create one consensus rating per NAEP standard with the help of a moderator and was not intended to allow for disagreements. This methodology was determined to be best suited to the scope and timing of this study. The consensus methodology is designed to highlight areas for states to examine, not to gather large amounts of data, record multiple ratings, or measure interrater reliability.

TABLE 1
Average ratings of alignment of Louisiana grade 5 benchmarks and grade level expectations and National Assessment of Educational Progress grade 4 science content statements

NAEP content area (number of NAEP standards)	Average rating
Overall physical science (15)	2.6
Matter (6)	2.7
Energy (5)	2.4
Motion (4)	2.8
Overall life science (7)	2.6
Structures and functions of living systems (4)	2.5
Changes in living systems (3)	2.7
Overall Earth and space science (11)	2.7
Earth and space in time (3)	3.0
Earth structures (3)	2.3
Earth systems (5)	2.8
All content (33)	2.6

Note: Rating is based on a scale of 1 to 3, where 1 indicates that state standards do not address the NAEP content statement, 2 that they partially address the NAEP content statement, and 3 that they fully address or exceed the NAEP content statement by the targeted grade level.

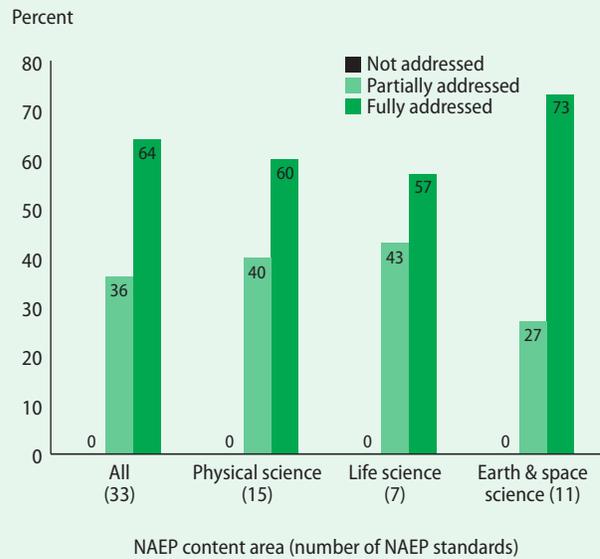
given ratings of 3, which means that state standards most often fully address the NAEP content statements (figure 1 and appendix C).

Reviewers observed that NAEP standards are based on descriptions of phenomena, whereas Louisiana’s standards are based on explanations of what a student does to explore the phenomena. In addition, reviewers found that NAEP standards tend to be widely inclusive (including, for example heat and electrical conductivity in the same standard) whereas most states tend to see such topics as part of two separate standards or benchmarks.

Areas of full alignment

Twenty-one NAEP grade 4 content statements are fully addressed by Louisiana benchmarks and grade level expectations. Nine of 15 physical science NAEP statements have full alignment with Louisiana, as do 4 of 7 life science statements and 8 of 11 Earth and space science statements.

FIGURE 1
The majority of Louisiana elementary school standards fully address National Assessment of Educational Progress content statements



Reviewers found Louisiana’s science standards to be rigorous and found consistent spiraling through the grade levels.

The 21 NAEP grade 4 content statements fully addressed by Louisiana are P4.1—measurable properties of objects and substances, P4.2—absorption, reflection, and conduction, P4.3—states of matter, P4.6—heating and cooling, P4.7—forms of energy (heat, electricity, light, sound), P4.10—vibrations and sound, P4.13—objects in motion; speed, P4.14 and P4.15—forces affecting motion, L4.3 and L4.4—interdependence of organisms, L4.5 and L4.6—heredity and reproduction, E4.1 and E4.2—objects in the universe and patterns in the sky, E4.3—history of Earth and Earth’s surface changes, E4.4—natural Earth materials, E4.7—role of the sun in Earth systems, E4.8 and E4.9—climate and weather, and E4.11—humans depend on and change their environments.

Areas of partial alignment

Thirty-six percent of the NAEP grade 4 content statements have partial alignment, in large part because many Louisiana standards imply content

explicitly stated by the NAEP and because NAEP content statements are often more detailed than Louisiana's.

Raters found that many Louisiana benchmarks imply content that the NAEP addresses in depth. For example, Louisiana PS-E-A5 and GLE PS 4.25 mention creating and separating mixtures, while NAEP's P4.4 covers content regarding objects being composed of single or multiple substances. The Louisiana standards likely imply the content regarding pure substances, but they do not explicitly state the single-substance composition of an object. In life science, the NAEP's L4.2 lists air, water, a source of energy, and light as needs for organisms, while LS-E-A1 and other corresponding grade level expectations do not specifically list such needs. Additionally, in Earth and space science, E4.10 lists fuels, metals, fresh water, and farmland as limited Earth resources, while Louisiana's matching benchmark and grade level expectations mention "resources" but do not specify those that the NAEP lists.

Reviewers also noted that teachers might more easily discern nuances in the standards if the grade level expectations, as well as the "key concepts" within the assessment guides, were integrated with the benchmarks.

The partially aligned NAEP standards are P4.4—composition of objects, P4.5—magnets, P4.8—heat, conductors, and increasing temperature, P4.9—light travel, P4.11—electrical circuits, P4.12—an object's position, L4.1—basic needs of organisms, L4.2—basic needs of plants and animals, L4.7—characteristics enabling survival in different environments, E4.5—properties of natural materials that sustain plant and animal life, E4.6—Earth materials for human use, and E4.10—limited supply of Earth resources.

Areas of nonalignment

There were no areas of nonalignment between NAEP grade 4 content statements and Louisiana grade 4 benchmarks.

Areas where Louisiana benchmarks go beyond the NAEP content statements

Louisiana has 60 benchmarks for grade 4. The NAEP does not address the 13 science as inquiry benchmarks, 5 of the 13 life science benchmarks, or 4 of the 13 Earth and space science benchmarks.

The NAEP addresses all of Louisiana's physical science and science and the environment benchmarks. However, the NAEP often implies content that is addressed in Louisiana, and in some instances, Louisiana contains more detailed physical science content than the NAEP. In science and the environment, the NAEP minimally and vaguely addresses all five Louisiana benchmarks.

The NAEP does not address the science as inquiry benchmarks because the NAEP discusses inquiry in a section separate from the content statements, called "science practices," intended to crosscut all NAEP content.

In life science, the NAEP does not address four of the six benchmarks in Louisiana's category of characteristics of organisms (LS-E-A2, A3, A5 and A6), which contains content regarding distinguishing between living and nonliving things; locating and comparing plant, animal, and human structures and functions; and recognizing the food groups necessary for maintaining a healthy body. The NAEP does not address three of the seven Earth and space science benchmarks, all in the category of properties of Earth materials (ESS-E-A5, A6 and A7), which contains content regarding the composition of rocks, variations in soil, and investigating fossils.

The majority of NAEP content statements for grade 4 are fully addressed by the Louisiana benchmarks and grade level expectations for grade 4

Summary of grade 4 alignment

All NAEP content items are at least partially addressed by Louisiana benchmarks and grade level expectations. The Louisiana benchmarks also

contain many items not listed in the NAEP's content statements. The combination of Louisiana's benchmarks and grade level expectations aligned very well with the NAEP content statements. The partially aligned content items were so because Louisiana's benchmarks and grade level expectations often implied content explicitly stated by the NAEP.

Reviewers made the general observation that NAEP standards are based on descriptions of phenomena, whereas Louisiana's standards are based on explanations of what a student does to explore the phenomena. They also found that NAEP standards tend to be widely inclusive (including, for example heat and electrical conductivity in the same standard), whereas most states see such topics as part of two separate standards or benchmarks. Reviewers found Louisiana's science standards to be rigorous and found consistent spiraling through the grade levels. However, they also noted that teachers might more easily discern nuances in the standards if the grade level expectations, as well as the "key concepts" in the assessment guides, were integrated with the benchmarks.

The majority of NAEP content statements are fully addressed by the Louisiana benchmarks and grade level expectations, and the overall alignment rating for Louisiana and NAEP science content at grade 4 is 2.6.

CONTENT ALIGNMENT AT GRADE 8

For grade 8, the NAEP provides 43 distinct content statements (displayed in parentheses in table 2). Seven (16 percent) of these content statements are fully addressed by Louisiana benchmarks, 33 (77 percent) are partially addressed, and three (7 percent) are unaddressed.

The average alignment rating for grade 8 is 2.1. The majority of content statements were given ratings of 2, which means that state standards partially address the NAEP content statements (figure 2 and appendix D).

TABLE 2

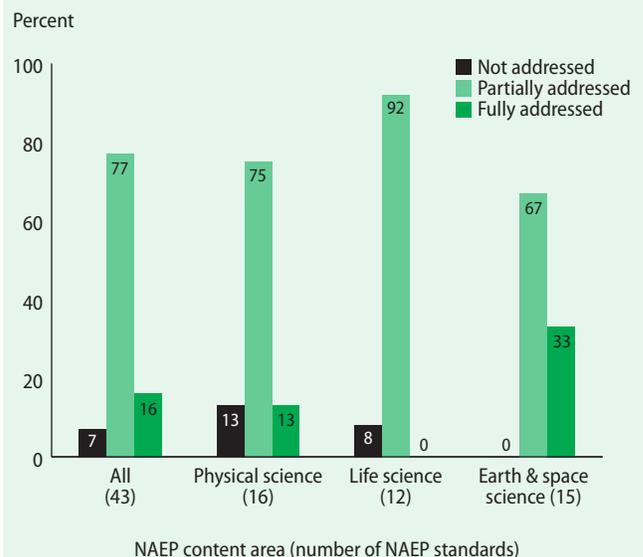
Average ratings of alignment of Louisiana middle school science benchmarks and grade level expectations and National Assessment of Educational Progress grade 8 science content statements

NAEP content area (number of NAEP standards)	Average rating
Overall physical science (16)	2.0
Matter (7)	2.0
Energy (6)	1.8
Motion (3)	2.3
Overall life science (12)	1.9
Structures and functions of living systems (8)	2.0
Changes in living systems (4)	1.8
Overall Earth and space science (15)	2.3
Earth and space in time (4)	2.5
Earth structures (6)	2.2
Earth systems (5)	2.4
All content (43)	2.1

Note: Rating is based on a scale of 1 to 3, where 1 indicates that state standards do not address the NAEP content statement, 2 that they partially address the NAEP content statement, and 3 that they fully address or exceed the NAEP content statement by the targeted grade level.

FIGURE 2

The majority of Louisiana middle school standards partially address National Assessment of Educational Progress content statements



Areas of full alignment

The seven NAEP content statements with which Louisiana is fully aligned are P8.1—properties of solids, liquids, and gases and the particulate model of matter, P8.16—forces, including magnitude, direction and net force, E8.2—gravity and the solar system, E8.4—Earth processes and the measurement of geologic time, E8.8—Earth’s layers, E8.12—causes of seasons, and E8.14—the water cycle. Five of the seven fully aligned content statements are in Earth and space science. Earth and space science is the section of Louisiana’s standards most closely aligned with the NAEP’s standards.

Areas of partial alignment

None of the 12 life science content statements is fully addressed by Louisiana benchmarks and grade level expectations. The vast majority of NAEP life science standards were given partial alignment ratings with Louisiana content, largely because many NAEP content statements are more detailed than the benchmark statements.

Reviewers often found that Louisiana’s standards do not address corresponding NAEP standards in the same amount of detail. For example, P8.5 deals with substances and their classifications according to physical and chemical properties. Louisiana’s matching standards (PS-M-A3 and GLE 6.4) cover grouping substances according to similar properties and differentiating between physical and chemical properties of substances. However, the NAEP’s standard also details classes of substances, such as metals and acids, while Louisiana does not provide such detailed examples. Another example is the alignment between the NAEP’s L8.1 and GLEs 5.15, 5.16 and 7.2. The NAEP contains content about the cellular composition of organisms, as well as the composition and functioning of cells, while the matching grade level expectations contain more general content about identifying and observing components and structures of cells.

Raters found that Louisiana’s benchmarks often imply content that the NAEP addresses in depth. For example, Louisiana’s LS-M-B1 refers to mitosis and meiosis, which may imply content regarding differentiation of cells and the formation of embryos, addressed in NAEP L8.2. In addition, E8.13 states that oceans affect climate because water holds a large amount of heat, while the Louisiana benchmark ESS-M-A11 and GLEs 8.27 and 8.44 do not go into similar depth regarding the heat-retaining properties of water and the ocean’s effect on weather.

Areas of nonalignment

Three NAEP content items are unaddressed by Louisiana benchmarks, two in physical science and one in life science. The items are P8.2—chemical properties of substances being explained by the arrangement of atoms and molecules, P8.13—nuclear reactions in the sun and photosynthesis, and L8.10—characteristics of organisms influenced by heredity and environment.

Louisiana’s benchmarks often emphasize different areas than does the NAEP. For example, the content area of heredity and reproduction is found in both the NAEP’s and Louisiana’s life science sections; however, the standards focus on different topics. Louisiana discusses chromosomes, genes, and passing traits onto offspring, while the NAEP details sexual and asexual reproduction, the importance of reproduction for the survival of a species, and heredity versus the environment.

Most of the Louisiana benchmarks and grade level expectations for grade 8 only partially address the NAEP content statements for grade 8

Areas where Louisiana benchmarks go beyond the NAEP content statements

Forty-seven Louisiana benchmark statements are not covered by the NAEP. The NAEP does not address the 15 Louisiana benchmarks in science as inquiry, 7 of the 22 in physical science, 9 of

the 16 in life science, 9 of the 23 in Earth and space science, or 7 of the 10 in science and the environment.

The NAEP does not address the 15 science as inquiry benchmarks because it discusses inquiry in a section separate from the content statements, called “science practices,” intended to crosscut all NAEP content.

In physical science, the NAEP does not address, for example, how factors such as temperature influence chemical reactions (PS-M-A8) or understanding that energy is involved in chemical reactions (PS-M-C7). In life science, the NAEP does not address human body systems and body changes (LS-M-A5 and A6) or investigating ecosystems (LS-M-C3). In Earth and space science, the NAEP does not address investigating characteristics of earthquakes and volcanoes (ESS-M-A3) or the causes and combating of coastal erosion (ESS-M-A8). The NAEP does not have a section devoted to science and the environment, so it does not address many of that section’s benchmarks, including distinguishing between renewable and nonrenewable resources (SE-M-A6) and identifying types of soil erosion and preventive measures (SE-M-A10). Some of the nonalignment between the NAEP and Louisiana’s science and the environment section may be due to the fact that the Louisiana section includes content on human and technological influence on the environment, while the NAEP covers similar ideas about the use of technological design in its “science practices” section.

Summary of grade 8 alignment

The Louisiana benchmarks cover many content topics that are not assessed by the NAEP, while most of the Louisiana content in its benchmarks and grade level expectations only partially addresses the NAEP content statements. This is primarily because many Louisiana standards imply NAEP content and because the NAEP is often more detailed in its presentation of content. The

overall alignment rating for grade 8 is 2.1, indicating partial alignment.

CONTENT ALIGNMENT AT GRADE 12

For grade 12, the NAEP provides 49 distinct content statements (displayed in parentheses in table 3). Twenty-five (51 percent) are fully addressed by Louisiana benchmarks, 23 (47 percent) are partially addressed, and 1 (2 percent) is not addressed.

The average alignment rating for grade 12 is 2.5. The majority of content statements were given ratings of 3, which means that most state standards fully address the NAEP content statements (figure 3 and appendix E).

Areas of full alignment

More than half the 49 NAEP grade 12 content statements are fully addressed by Louisiana benchmarks and grade level expectations. Sixty-five percent of all physical science content statements are fully addressed, as are 46 percent of life science statements and 31 percent of Earth and space science statements. The level of detail of the Louisiana grade level expectations often matches that of the NAEP standards, and the combination of benchmarks and grade level expectations often warranted a rating of full alignment with a very detailed NAEP statement.

Areas of partial alignment

Twenty-three NAEP grade 12 content statements (47 percent) have partial alignment, largely because Louisiana benchmarks often imply content explicitly stated by the NAEP and because NAEP content statements are often more detailed than Louisiana’s.

Each NAEP content statement is matched to an average of five Louisiana content statements (benchmarks or grade level expectations), since each NAEP standard is highly detailed and incorporates ideas that are put into separate benchmarks and

Most Louisiana standards for grade 11 fully address the NAEP content statements for grade 12

TABLE 3

Average ratings of alignment of Louisiana high school benchmarks and grade level expectations and National Assessment of Educational Progress grade 12 science content statements

NAEP content area (number of NAEP standards)	Average rating
Overall physical science (23)	2.6
Matter (7)	2.7
Energy (9)	2.4
Motion (7)	2.7
Overall life science (13)	2.5
Structures and functions of living systems (7)	2.1
Changes in living systems (6)	2.8
Overall Earth and space science (13)	2.3
Earth and space in time (7)	2.3
Earth structures (1)	2.0
Earth systems (5)	2.4
All content (49)	2.5

Note: Rating is based on a scale of 1 to 3, where 1 indicates that state standards do not address the NAEP content statement, 2 that they partially address the NAEP content statement, and 3 that they fully address or exceed the NAEP content statement by the targeted grade level.

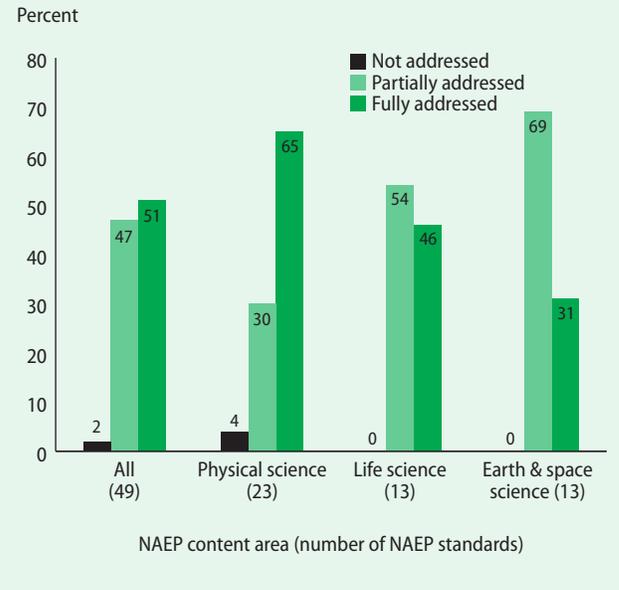
grade level expectations by Louisiana. The majority of partially aligned statements were coded “IC,” denoting implied content in Louisiana’s standards.

NAEP content statements at grade 12 are characteristically detailed and extensive. Each content statement typically contains several sentences, each of which can often be matched to a single Louisiana benchmark. Therefore, much of the alignment between the NAEP grade 12 and Louisiana’s grade 11 benchmarks consists of several Louisiana benchmarks and grade level expectations matched and rated with a single NAEP statement.

Raters found that many Louisiana benchmarks imply content that the NAEP addresses in depth. For example, Louisiana’s PS-H-B2 describes the nature and importance of radioactive isotopes and nuclear reactions (fission, fusion, radioactive decay), while the NAEP’s P12.11 defines fission and fusion, which presumably is implied in descriptions of their nature and importance.

FIGURE 3

The majority of Louisiana high school standards fully address National Assessment of Educational Progress content statements



Part of L12.9 was matched with LS-H- B3 but was given an implied content code because mention of altered genes is found in the NAEP but not in Louisiana. E12.6 has content regarding early Earth, including evidence for bacteria and the absence of atmospheric oxygen, while SE-H-A7 and A8 and GLE HS EnvSci.9 deal generally with the biosphere and the evolution and adaptation of plants and animals but do not provide such detailed illustrations of early Earth.

Raters also found that the NAEP’s content often provides more detail than do the Louisiana benchmarks. For example, while the NAEP’s E12.7 names some of the geological processes of Earth, including earthquakes, volcanic eruptions, the building of mountain chains, and the shifting of continents, Louisiana’s matching standards (SE-H-A6, SE-H-A4, ESS-H-C5, GLE EnvSci.5, GLE EarthSci.22) mention changes in Earth’s structure more generally.

Areas of nonalignment

One NAEP content item is unaddressed by Louisiana benchmarks. That item is P12.13, which states,

“The potential energy of an object on Earth’s surface is increased when the object’s position is changed from one closer to Earth’s surface to one farther from Earth’s surface.”

Areas where Louisiana benchmarks go beyond the NAEP content statements

More than one-third of Louisiana benchmarks in grade 11 are not covered by the NAEP. The NAEP does not address the 12 Louisiana benchmarks in science as inquiry, 8 of the 29 in physical science, 13 of the 30 in life science, 3 of the 18 in Earth and space science that are used in the state assessments, and 2 of the 12 in science and the environment that are used in the state assessments.

The Louisiana benchmarks for grade 11 contain many items that are not assessed by the NAEP

The NAEP does not address the 12 science as inquiry benchmarks because it discusses inquiry in a section separate from the content statements, called “science practices,” intended to crosscut all NAEP content. In

physical science, the NAEP does not address, for example, several Louisiana benchmarks about chemical reactions (PS-H-D1, D2, D3 and D4). In life science, the NAEP does not address several benchmarks in the category of biological evolution (LS-H-C5, C6 and C7), nor does it address any of the benchmarks in the categories of systems and the behavior of organisms (LS-H-F1, F2, F3, and F4) or personal and community health (LS-H-G1, G2, G3, G4 and G5). The NAEP does not address Louisiana’s highly specific Earth and space science benchmark about the geologic development of Louisiana (ESS-H-C3), nor does it address the demonstration of the laws of motion for orbiting bodies (ESS-H-D6). ESS-H-D7 and two science and the environment benchmarks (SE-H-B6 and SE-H-C4) involve human and technological influence on the environment and are not covered by NAEP content statements, which cover similar ideas regarding the use of technological design in the “science practices” section.

Summary of NAEP grade 12 alignment

The Louisiana benchmarks contain many items that are not assessed by the NAEP. Often, Louisiana benchmarks were noted for their alignment to distinct parts of a more detailed NAEP content statement. Overall, Louisiana is well aligned to the NAEP. Many Louisiana high school standards only partially address the NAEP content statements at grade 12 because they imply NAEP content and because the NAEP is often more detailed and in-depth in its presentation of content. However, there is a generally high alignment between NAEP and Louisiana, and the overall alignment rating is 2.5.

TEST SPECIFICATIONS ALIGNMENT

The assessment specifications alignment involved two parts: examining the types of items found in the NAEP and in the Louisiana Educational Assessment Program and the Graduation Exit Examination, and comparing the NAEP’s distribution of items among the different science strands with that of the Louisiana Educational Assessment Program and the Graduation Exit Examination.

Science is a discipline with a strong tradition of investigation, experimentation, and application of knowledge and skills. Before the 2005 assessment, NAEP science assessments consisted primarily of short-answer, paper-and-pencil questions that were mostly multiple-choice, which can go only so far in assessing skills. To improve the assessment of the range of science knowledge and skills, the last two NAEP science frameworks have expanded the range of item types on the test. In particular, the 2009 NAEP framework takes advantage of advances in educational measurement and the development of computer-based assessments. Due to the varying ways that differing item types assess and reveal what students know and can do, the NAEP 2009 assessment specifications require future NAEP tests to incorporate a range of item types, allowing students to reveal their understanding in ways beyond traditional selected-

response methods. Multiple-choice items, short constructed-response items, extended constructed-response items, hands-on performance tasks, and interactive computer tasks will all be used to more accurately assess student knowledge, thinking, and skills.

Each type of assessment item demands a unique response from students (selecting a response from a set of alternatives, writing an explanation or justification, performing a virtual lab experiment). Individual items may draw on different types of stimuli (verbal, graphic, manipulative) to access the knowledge and skills required or may be scored in a variety of ways (right/wrong, partial credit, human scorers, computer software). By using several types of items the 2009 NAEP science assessment will require students to draw on multiple types of knowledge and a variety of skills for using and expressing that knowledge, thereby giving a more accurate picture of the breadth and depth of their learning. In this study, the following item types from the NAEP were compared with the types in use by the states.

In multiple-choice items, students reflect on the material and then select an answer from a limited number of alternatives. Well constructed multiple-choice items can probe important facts, broad concepts, and themes of science, as well as deductive reasoning skills.

Constructed-response items, in which students answer without reference to a provided list of alternatives, include short constructed-response items and extended constructed-response items. Constructed-response items can provide insight into students' levels of conceptual understanding and assess their ability to communicate about science. They can also be used to probe student ability to generate information related to science content statements and their interconnections (how two or more cyclic events are related). Constructed-response items may be particularly useful for probing the practices of using scientific inquiry or using technological design (interpret given data or provide a solution to a real-world problem).

In hands-on performance tasks, students manipulate selected physical objects and try to solve a scientific problem involving the objects. These exercises, if carefully designed, can probe student abilities to combine science knowledge with the investigative skills reflective of the nature of science and inquiry.

Interactive computer tasks in the 2009 NAEP science assessment may involve information search and analysis, empirical investigation, simulation, or concept mapping. The broad purpose of interactive computer tasks in this context is to tap

performance expectations that are more advantageously assessed in a virtual format, such as scientific modeling of microscopic or temporal phenomena, repeated experiments, or simulations of hazardous or messy lab situations. Interactive computer tasks are intended as a complement to the hands-on performance tasks, not as a replacement.

The NAEP specifications also include two other types of items, item clusters and predict-observe-explain item sets. Item clusters are groups of related items that provide more in-depth analysis of student performance than would a collection of discrete, unrelated items. They can be particularly useful in exploring student conceptions, predictions, or explanations of the natural world. The predict-observe-explain item sets (White & Gunstone, 1992) describe a situation and ask the student to predict, observe, or explain the outcome, sometimes with additional supporting detail. Predict-observe-explain items may involve using science principles or the cognitive demand of “knowing why (schematic knowledge).” Because these are really ways of clustering items and are not usually included in state test specifications, they were not used for comparison in this study.

By using several types of items the 2009 NAEP science assessment will require students to draw on multiple types of knowledge and a variety of skills for using and expressing that knowledge

The NAEP stipulates that 50 percent of student response time should be spent on multiple-choice items and the other 50 percent on constructed-response items (including short constructed-response, extended-constructed-response, and concept-mapping tasks). Within these two categories are item clusters, predict-observe-explain item sets, hands-on performance tasks, and interactive computer tasks. There will be at least one item cluster, one predict-observe-explain item set, one hands-on performance task, and one interactive computer task at each grade level, and the total number of interactive computer tasks plus hands-on performance tasks will be at least four at each grade level.

The number of score points in the Louisiana tests is the same for grades 4, 8, and 11 (grade 11 is compared to the NAEP grade 12 in this analysis). There are 40 multiple-choice items, 7 short constructed-response items, and 1 extended constructed-response item at each grade. Three of the short answer items and the extended constructed-response item are part of a “comprehensive science task” that is also given at each grade level. Of the combined 47 multiple-choice and short constructed-response items, approximately two-thirds address the NAEP content strands of physical, life, and Earth and space sciences. The remaining items address other science content from the state benchmarks.

Table 4 shows the percentages of various item types found in the NAEP and in Louisiana. The

2009 NAEP will have 50 percent of student response time allocated to multiple-choice items and 50 percent allocated to constructed-response items (short and extended). The current Louisiana tests do not have their item distributions proportioned by student response time, so the table shows the NAEP’s proportions of student response time and Louisiana’s proportions of items and points. The proportions of items and points are the same for Louisiana grades 4, 8, and 11. The Louisiana test contains mostly multiple-choice items, but it does include short constructed-response items, an extended constructed-response item, and a “comprehensive science task.” The distribution of points for Louisiana at all grade levels is fairly similar to the distribution of time for the NAEP; Louisiana is 69 percent multiple-choice and 31 percent constructed-response, while the NAEP is 50 percent multiple-choice and 50 percent constructed-response. Louisiana’s test specifications indicate that its tests will include “comprehensive science tasks,” which require students to read, use, and react to a scenario that typically includes diagrams, data tables, and graphs, and may require students to complete or interpret data tables or to record observations. This is different from the NAEP hands-on performance task, which requires the manipulation of physical objects in order to solve a scientific problem.

To consider how the state test coverage of the NAEP science topics matched, table 5 shows the proportions of testing time devoted to each of the three content areas for the NAEP and for the

TABLE 4

Proportions of different item types on the Louisiana science assessment (percent)

NAEP item types	NAEP	Louisiana (grades 4, 8, 11)	
	Share of time	Share of total items	Share of total points
Multiple-choice items	50	83	69
Short constructed-response items	50	15	24
Extended constructed-response items		2	7
Hands-on performance tasks ^a	(≥1)		
Interactive computer tasks ^a	(≥1)		

a. Hands-on performance tasks and interactive computer tasks are combination items and can be categorized as multiple-choice or constructed-response.

Louisiana test. The first column of the table lists all the science topic areas that are included on the Louisiana test. The first three topic areas (physical, life, and Earth and space sciences) are those that are covered on the NAEP, while the two topics below those (science as inquiry and science and the environment) are not separately assessed on the NAEP.

Under the column heading for grade 4, three sub-columns are shown. The first shows the proportion of testing time devoted to each of the three NAEP topic areas. The second shows the proportion of points Louisiana allots to each strand, excluding the comprehensive science task, which includes three short-answer questions about science as inquiry and one extended constructed-response item about one of the four content strands (in grades 4 and 8) or two of the four content strands (in grade 11). (The comprehensive science task is excluded from these calculations because it is unknown which of the four content strands it will contain.) The third subcolumn shows the

comparison of the proportions devoted to the three NAEP topics, a positive number if the Louisiana test devotes more and a negative number if the NAEP devotes more. This pattern of columns is repeated for middle and high school. Louisiana grade 11 was compared with the NAEP grade 12.

At all grade levels, the proportion of points devoted by Louisiana to each NAEP content strand is less than the proportion of time devoted to each strand by the NAEP. This is because Louisiana allots a significant proportion of its test to science as inquiry and science and the environment, which are not separate strands on the NAEP.

Table 6 ignores the testing time devoted to science as inquiry and science and the environment, which are not separately tested in the NAEP, and shows how the proportions of testing time on the NAEP, for the three NAEP strands, compare with the proportions of points on the state test. These calculations exclude the comprehensive science task because it is not known which of the strands

TABLE 5

Approximate testing time allocated to different science topics on the Louisiana science assessment (percent)

Content area	Grade 4			Grade 8			Grade 12/Grade 11		
	NAEP (time)	Louisiana (points)	Difference	NAEP (time)	Louisiana (points)	Difference	NAEP (time)	Louisiana (points)	Difference
Physical science	33	21	-12	30	21	-9	37.5	25	-13
Life science	33	21	-12	30	21	-9	37.5	25	-13
Earth and space science	33	21	-12	40	21	-19	25	17	-8
Science as inquiry	0	17		0	17		0	17	
Science and the environment	0	21		0	21		0	17	

TABLE 6

Comparison of the proportions of testing time allocated to the National Assessment of Educational Progress science topics (percent)

Content area	Grade 4			Grade 8			Grade 12/Grade 11		
	NAEP (time)	Louisiana (points)	Difference	NAEP (time)	Louisiana (points)	Difference	NAEP (time)	Louisiana (points)	Difference
Physical science	33	33	0	30	33	+3	37.5	37.5	0
Life science	33	33	0	30	33	+3	37.5	37.5	0
Earth and space science	33	33	0	40	33	-7	25	25	0

it will contain. At the elementary level, the proportions are the same in the NAEP and in Louisiana. In grade 8, Louisiana devotes 33 percent of its points equally to all three topics, whereas the NAEP emphasizes Earth and space science by allocating 40 percent of testing time to it and only

30 percent to the other two topics. For grade 12, the NAEP devotes more time to the physical and life science strands, and Louisiana also gives more points to the physical and life science strands. The Louisiana proportions match the NAEP proportions almost exactly.